

1. Üzenet és entrópia

Jelkészlet: a,b,c,d,e,f,g,h (8 db jel)

Forrásüzenet: fdbeccfbfdebfbafdfbhhahgffebdeagdcabacachcbffhfhfd (50 karakter)

Shannon egy véges sok jeltől álló (véges ábécé feletti) üzenet információértékét az üzenet jeleinek mint a jelle jellemző valószínűséggel bekövetkező események információtartalmának „átlagos”, azaz várható értékeként

$$H(x_1, x_2, \dots, x_j) = \sum_{i=1}^j p(x_i) I(x_i)$$

határozta meg: azaz az i -edik jelhez tartozó információ mennyiségével kifejezve az entrópia az a jelnekénti információ valószínűség és az elem gyakoriságának szorzatának elemenként

vett összege $\sum_{i=1}^n p(x_i) \log_2 \frac{1}{p(x_i)}$. Az entrópia egy elem bizonytalanságának mértékével egyenlő. "Ha két jel azonos értékkel rendelkezik akkor az azt jelzi, hogy nehéz megjósolni a következő elemet."

$$H(x) = 0,12 * 3,058893689 + 0,16 * 2,64385619 + 0,12 * 3,058893689 + 0,12 * 3,058893689 + 0,08 * 3,64385619 + 0,22 * 2,184424571 + 0,04 * 4,64385619 + 0,14 * 2,836501268 = 2,879165$$

Az entrópia: 2,879165

jel	valószínűség	kódhossz 1	kódhossz 2	kódhossz 3	kódhossz 4	(-)log2 és szám	* valószínűség	kód * valószínűség 1	kód * valószínűség 2
a	0,12	3	1	3	3	3,058893689	0,367067243	0,36	0,12
b	0,16	3	1	3	3	2,64385619	0,42301699	0,48	0,16
c	0,12	3	1	3	3	3,058893689	0,367067243	0,36	0,12
d	0,12	3	1	3	3	3,058893689	0,367067243	0,36	0,12
e	0,08	3	1	3	3	3,64385619	0,291508495	0,24	0,08
f	0,22	3	1	3	3	2,184424571	0,480573406	0,66	0,22
g	0,04	3	1	3	3	4,64385619	0,185754248	0,12	0,04
h	0,14	3	1	3	3	2,836501268	0,397110177	0,42	0,14
Forráskódhossz:	50					Entrópia:	2,879165044	3	1

2. Szinguláris

Szinguláris kód, mert minden abc-hez ugyanazt rendeljük.

a – 0

b – 0

c – 0

d – 0

e – 0

f – 0

g – 0

h – 0

Átlagos kódhossz: (A szinguláris kódhossz, jelen esetben:) 1

3. Azonos hosszúságú

a – 000

b – 001

c – 011

d – 010

e – 100

f – 101

g – 110

h – 111

Az üzenet kódolva:

101010001100011011101001101010100001101001000101010101001111111000111110101
10110000101010000011001001100000100001100001111101100110110111111111101010

Átlag kódhossz: 3

Mindig készíthetünk olyan prefix kódot, amelynek átlagos kódhossza kisebb mint $m(\min_i p_i) = \frac{H(X)}{\log_2 r} + 1$. Következésképp ennek a kódnak valamennyi kódszava m -nél rövidebb kell, hogy legyen. Vagyis ez is a fenti osztályba tartozik. Ezért az $L_{\min}(X)$ minimális átlagos kódhosszra fennáll, hogy

$$\left(\frac{H(X)}{\log_2 r} \leq \right) L_{\min}(X) < \frac{H(X)}{\log_2 r} + 1.$$

Ezzel állításunkat igazoltuk. \square

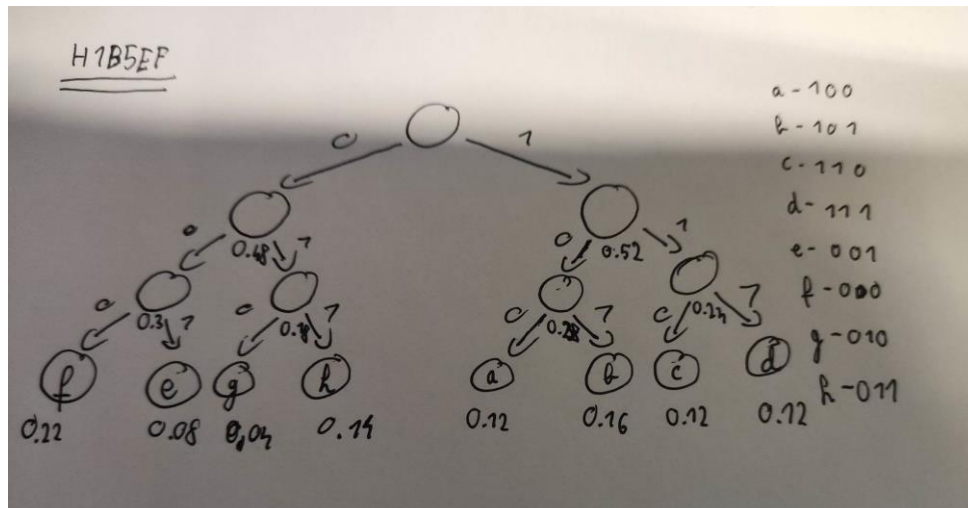
A maximális hatásfokú egyértelműen dekódolható kódokat *optimális* ilyen több is lehet, de biztosan van köztük prefix (irreducibilis) kód. ezeket vizsgáljuk.

Minimális átlagos kódhossz: $2,879165 < 3 < 3,879165$

4. Változó hosszúságú

Az üzenet kódolva:

000111101001110110000101000111001101000101100000111000101011011100011010000
000001101111001100010111110100101100110100110011110101000000011011011000111



Minimális átlagos kódhossz: $2,879165 < 3 < 3,879165$

Átlag kódhossz: 3